Ocean Modelling 81 (2014) 25–35

Contents lists available at ScienceDirect

Ocean Modelling

journal homepage: www.elsevier.com/locate/ocemod

Assessing the utility of frequency dependent nudging for reducing biases in biogeochemical models

Karl B. Lagman, Katja Fennel *, Keith R. Thompson, Laura Bianucci

Department of Oceanography, Dalhousie University, PO Box 15000, Halifax, NS B3H 4R2, Canada

ARTICLE INFO

Article history:
Received 31 October 2013
Received in revised form 16 June 2014
Accepted 25 June 2014
Available online 10 July 2014

Keywords:
Bias reduction
Numerical modeling
Biogeochemical modeling
North Atlantic

ABSTRACT

Bias errors, resulting from inaccurate boundary and forcing conditions, incorrect model parameterization, etc. are a common problem in environmental models including biogeochemical ocean models. While it is important to correct bias errors wherever possible, it is unlikely that any environmental model will ever be entirely free of such errors. Hence, methods for bias reduction are necessary. A widely used technique for online bias reduction is nudging, where simulated fields are continuously forced toward observations or a climatology. Nudging is robust and easy to implement, but suppresses high-frequency variability and introduces artificial phase shifts. As a solution to this problem Thompson et al. (2006) introduced frequency dependent nudging where nudging occurs only in prescribed frequency bands, typically centered on the mean and the annual cycle. They showed this method to be effective for eddy resolving ocean circulation models. Here we add a stability term to the previous form of frequency dependent nudging which makes the method more robust for non-linear biological models. Then we assess the utility of frequency dependent nudging for biological models by first applying the method to a simple predator–prey model and then to a 1D ocean biogeochemical model. In both cases we only nudge in two frequency bands centered on the mean and the annual cycle, and then assess how well the variability in higher frequency bands is recovered. We evaluate the effectiveness of frequency dependent nudging in comparison to conventional nudging and find significant improvements with the former.

1. Introduction

The bias of an estimator is formally defined as the difference between its expected value and the true value it is trying to estimate (e.g., Priestley, 1981). In the context of environmental modeling, biases are often approximated by the mean difference between simulated and observed quantities after averaging over certain temporal or spatial scales (e.g., WMO, 2008). Biases are a common problem in many environmental models (e.g., Randall et al., 2007) and are sometimes considered severe enough to preclude the direct use of model output. For example in atmospheric studies of climate change impacts, bias reduction is a standard procedure (see Ehret et al., 2012 and references therein). The averaging time-scale for bias calculation can range from a few days for the verification of synoptic forecasts to decades for the verification of climate models. Observational climatologies are often used to calculate biases over seasonal and longer time-scales. Biases can be caused by many factors including incorrect model parameterizations, insufficient model resolution, discretization errors, incorrect or imperfect open boundary conditions and forcing, and are to be expected in most models of the natural world.

Model drift and the associated biases are a common problem with biogeochemical ocean models (e.g., Nerger and Gregg, 2007; Doney et al., 2009; Lehmann et al., 2009; While et al., 2010). Errors in biological variables can be inherited from problems in model physics, e.g. subtle biases in vertical mixing that do not lead to obvious problems in physical fields but can result in notable errors in phytoplankton concentrations because the latter are highly sensitive to vertical nutrient supply. Biases can also result from problems with the biogeochemical model itself, e.g. incorrect process resolution or imperfect parameterizations.

It is important not only to quantify biases but also to understand their causes and correct them where possible. Diagnosing bias errors can elucidate systematic problems in model formulation such as incorrect parameterizations and ultimately lead to improved models. However, it is unlikely that any deterministic model will ever be completely free of these errors, hence techniques for bias reduction are necessary. Moreover, many sequential data assimilation techniques (e.g. Kalman Filters) assume bias-free observations and model states. When applying these methods, biases should be removed first. It has been shown that

* Corresponding author.
E-mail address: katja.fennel@dal.ca (K. Fennel).

http://dx.doi.org/10.1016/j.ocemod.2014.06.006
1463-5003/

bias reduction improves the results of data assimilation in atmospheric applications (Dee and Todling, 2000; Baek et al., 2006), physical ocean models (Chepurin et al., 2005; Keppenne et al., 2005) and ocean biogeochemical models (Nerger and Gregg, 2008; While et al., 2010).

Bias has long been recognized as a serious problem in atmospheric and ocean modeling (e.g., Doney et al., 2009) and various suppression techniques have been developed. For example, offline bias reduction during post-processing of model output is a standard tool in atmospheric modeling (Ehret et al., 2012). Perhaps the simplest method for online bias reduction is nudging, where simulated fields are continuously forced toward direct observations or a climatology. During each time step an increment proportional to the difference between observation and model is scaled by an inverse relaxation time and added to the field being corrected. Henceforth we will refer to this method as *conventional nudging*.

Conventional nudging is widely used in biogeochemical ocean models. In fact, early biogeochemical models relied on nudging (then also referred to as restoring) of model nutrients to climatological nutrient distributions in order to infer net community production and other biogeochemical processes (Najjar et al., 1992; Marchal et al., 1998). In the more recent, mechanistically detailed biogeochemical models nudging is frequently used for the reduction of biases resulting from imperfect boundary conditions; for instance, in nested 3D applications variables are nudged to physical and biogeochemical distributions (either from lower-resolution, larger-scale models or climatological observations) in buffer zones along their open boundaries (e.g. Fennel et al., 2008). Nudging is also used in 1D models to drive variables toward either direct observations (e.g. Bagniewski et al., 2011) or climatologies (e.g. Fennel et al., 2003) in order to account for unresolved 3D processes.

Advantages of conventional nudging are that it is easy to implement, robust and can force the model arbitrarily close to the observations. Unfortunately, there are serious limitations as well if the technique is used to nudge a model towards a climatology: high-frequency variability (e.g., eddies in ocean circulation models) are suppressed and artificial phase lags are introduced, especially when nudging is strong (Woodgate and Killworth, 1997; Thompson et al., 2006). As a solution to this problem, Thompson et al. (2006) proposed limiting the nudging to prescribed frequency bands, leaving the model to evolve freely outside of these bands. We will refer to this modified method as *frequency dependent* nudging. (In the original paper by Thompson et al. (2006) the nudges were filtered in both space and time and, for this reason, the original method was called spectral nudging. In the present application the nudges are only filtered in time (i.e., in the frequency domain) and so we will refer to the method as frequency dependent nudging.) In ocean and atmosphere models the chosen frequency bands are often centered on the mean and annual cycle, which tend to be well characterized in climatologies. Frequency dependent nudging has been demonstrated to be effective in reducing bias errors in eddy resolving ocean circulation models (e.g. Thompson et al., 2006, 2007; Stacey et al., 2006; Zhu et al., 2010).

Here we perform an exploratory study to assess the utility of frequency dependent nudging in reducing seasonal biases in biogeochemical ocean models without suppressing higher frequency variations (e.g., blooms with typical scales of a week). To our knowledge, frequency dependent nudging has not yet been applied to such models. We use a framework where a simulation from a *complete model* is sampled and these samples are treated as *observations*. Although these "observations" are synthetic we henceforth refer to them simply as observations. A *climatology*, defined to consist only of the mean and annual cycle (i.e. a sinusoid with a period of 1 year), is then derived from these observations. A *simplified model* (which has been degraded compared to the complete model in terms of its functionality) is run without any bias reduction. In the cases studied here, the simplified model results in significant bias errors. Conventional and frequency dependent nudging of the simplified model toward the climatology defined by the mean and annual cycle are then used to suppress biases in the model states. By comparing the nudged simulations against the observations, we assess to what degree the nonlinearity of the models is able to recover the true variability in the higher frequency bands for both nudging schemes. We present two examples: The first is a simple predator–prey model; the second is a 1D biological model configured for the shelf seas in the northwestern North Atlantic. Taken together, these two examples can be seen as a first step toward applying frequency dependent nudging to a more complete, 3D biogeochemical model of the region.

The structure of the paper is as follows. An overview of frequency dependent nudging is provided in Section 2. We apply our framework to the simple predator–prey model in Section 3 and illustrate the potential benefits and problems of frequency dependent compared to conventional nudging. In Section 4 we apply the same steps to a 1D biological ocean model, followed by a summary in Section 5.

2. Overview of nudging

To motivate the form of frequency dependent nudging used here, and illustrate how it differs from conventional nudging, consider the following linear equation for the evolution of the p-dimensional state vector x:

$$\frac{dx}{dt} = \Phi x + f \tag{1}$$

where Φ is a time-invariant system matrix and f represents the time-dependent forcing. The real parts of the eigenvalues of Φ are assumed negative thereby ensuring asymptotic stability.

A simple way to reduce the discrepancy between the model state and an observed climatology, $c(t)$, is to add a simple conventional nudging term of the form $\gamma(c - x)$ to Eq. (1):

$$\frac{dx}{dt} = \Phi x + f + \gamma(c - x) \tag{2}$$

where γ is the nudging coefficient. Note that conventional nudging does not alter the stability of the model because the real parts of the eigenvalues of the modified dynamics matrix $(\Phi - \gamma I)$ do not change sign. If $\gamma = 0$ then $x(t)$ will equal the un-nudged state. As $\gamma \to \infty, x(t)$ will tend toward the climatology to which the model is nudged. For simplicity, we have assumed the climatology is available for every element of the state vector.

Fourier transforming Eq. (2) at frequency ω leads to

$$X_n = (i\omega I - \Phi + \gamma I)^{-1}[(i\omega I - \Phi)X_u + \gamma C] \tag{3}$$

where $X_n(\omega)$ and $C(\omega)$ are the Fourier transforms of the nudged state and climatology, respectively. $X_u(\omega)$ is the Fourier transform of the un-nudged state and is obtained by Fourier transforming Eq. (1).

It is clear from Eq. (3) that X_n is a weighted average of X_u and C. As $\gamma \to \infty$, the difference between the observed and simulated climatology will tend to zero but the variability at non-climatological frequencies (where $C = 0$) will be attenuated. To solve these problems, Thompson et al. (2006) suggested that the nudging be limited to frequency bands centered on climatologically relevant frequencies (e.g., 0 and 1 cycle per year); outside of these frequency bands the model is not nudged and can evolve freely. This corresponds to replacing (2) by

$$\frac{dx}{dt} = \Phi x + f + \gamma\langle c - x\rangle \tag{4}$$

where $\langle\cdot\rangle$ denotes a quantity that has been bandpass filtered to pass variations in the vicinity of climatologically relevant frequencies.

If γ is sufficiently small that Eq. (4) remains stable, the Fourier transform of the nudged state is still given by (3) if we replace γ by $\gamma\Gamma(\omega)$ where $\Gamma(\omega)$ is the transfer function of the bandpass filter. It follows that, away from the climatological frequencies where $\Gamma(\omega) = 0, X(\omega) = X_u(\omega)$ as expected.

Biogeochemical models are highly nonlinear and so we now generalize Eq. (1) to

$$\frac{dx}{dt} = \phi(x, t) \tag{5}$$

where the dependence of ϕ on time t allows for the possibility of time-dependent parameters and external forcing. Based on the above discussion we propose the following form of frequency dependent nudging:

$$\frac{dx}{dt} = \phi(x, t) + \gamma\langle c - x\rangle + \delta(c - x) \tag{6}$$

Note that this equation differs from Eq. (2) through the addition of a conventional nudging term with nudging coefficient δ. This term was added to increase the stability of the nudged system. Details on the implementation of the bandpass filter are given below.

Biogeochemical models can generate, and couple, variability across a wide range of time scales. Hence, it is not clear a priori that the frequency dependent nudging defined in Eq. (6) will work nor that it will work better than conventional nudging. In the next section the effectiveness of the scheme is evaluated using one of the simplest models of predator–prey interactions: a modified Lotka–Volterra model.

3. Nudging a simple predator–prey model

A highly idealized model of the interaction of prey (x_1) and predators (x_2) is

$$\begin{aligned}\frac{dx_1}{dt} &= \alpha_1 x_1(1 - x_1/\alpha_3) - \alpha_4 x_1 x_2\\ \frac{dx_2}{dt} &= \alpha_5 x_1 x_2(1 - x_2/\alpha_6) - \alpha_2 x_2\end{aligned} \tag{7}$$

where α_1 and α_3 control the growth of the prey and α_4 controls the rate of predation, α_5 and α_6 control the growth of the predators and α_2 is their mortality rate.

This pair of equations differs from the well known Lotka–Volterra model in one important respect: the growth terms for prey and predators use the logistic growth parameterization instead of a constant growth rate. The constant growth rates in the standard Lotka–Volterra equations assume infinite carrying capacities. The above modification addresses this issue, implicitly representing resources via an imposed carrying capacity for both prey and predators. The carrying capacities for prey and predators are α_3 and α_6, respectively. Modified Lotka–Volterra (LV) equations such as Eq. (7) have been discussed extensively in the ecological literature (e.g. MacArthur, 1970; May, 1973; Chesson, 1990; Berryman, 1992).

To simplify Eq. (7) we assume $\alpha_1 = \alpha_2 = \alpha_4 = \alpha_5 = 50$ and $\alpha_3 = \alpha_6 = 2$. For these parameters the model (LV0) has a fixed point at $(1.236,\ 0.382)$. Any trajectory that starts in the vicinity of this point will spiral inwards with an e-folding time of 0.0468. An example of such a trajectory is shown by the gray line in Fig. 1.

Next, we allow the carrying capacity of the prey to vary with time (LV1) as follows

$$\alpha_3 = \alpha_{30}[1 + \alpha_{31}\sin(2\pi t) + \alpha_{32}\sin(2\pi t/P_2)]. \tag{8}$$

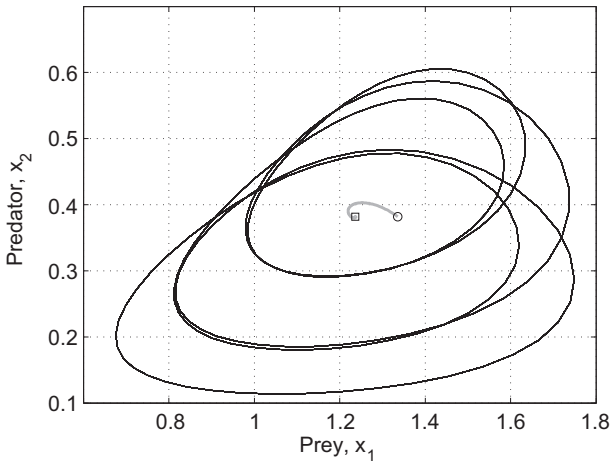

Fig. 1. Evolution of prey and predator abundance according to Eq. (7). The grey trajectory was calculated using a constant α_3 (run LV0). The circle indicates the initial condition and the square indicates the interior fixed point. The black line was calculated using a time-varying α_3 (run LV1). The trajectory effectively reaches a periodic steady state after a spin up of one year. The plotted trajectory is from a 40 year run excluding the first year. The parameter values for both runs are given in Table 1.

We interpret the term $\sin(2\pi t)$ as a variation of the carrying capacity with a period of one year, and $\sin(2\pi t/P_2)$ as a high frequency variation about this annual cycle with a period P_2 years. We assume $P_2 = 0.2$ years.

The impact of allowing α_3 to vary with time is shown by the black lines in Figs. 1 and 2. (Parameter values for this run are given in Table 1.) As expected, the prey and predator abundances now vary with periods of 1 and 0.2 years. The nonlinearity of the governing equations also generates variability at other periods. This can be seen in the way the prey abundance varies with greater amplitude at about the annual cycle when the predator abundance is low (e.g., $39.5 < t < 40$). The impact of turning off the annual cycle ($\alpha_{31} = 0$) is shown by the gray lines in the left panels of Fig. 2. As expected, the variations of x_1 and x_2 are almost perfectly sinusoidal with a period of 0.2 and no evidence of an annual cycle.

We now perform a set of numerical experiments to compare the effectiveness of conventional and frequency dependent nudging in reducing seasonal biases in the model state. All of the model runs

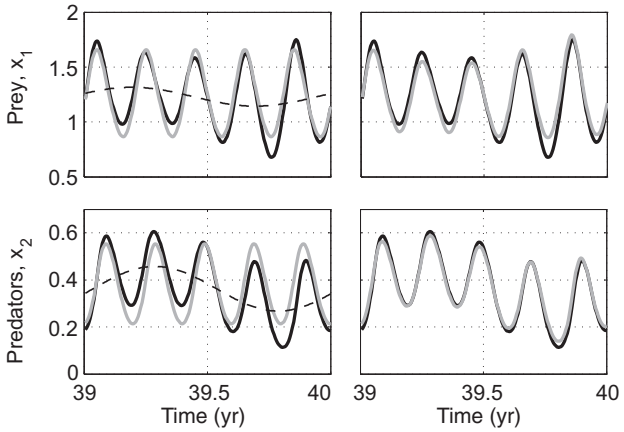

Fig. 2. Time variation of prey and predator abundance according to Eq. (7). The time axis shows the last annual cycle of a 40 year run. The black lines show the response to the full time variation of α_3 (run LV1). The dashed lines show the climatology of this run consisting of the mean and a sinusoid with period 1 year. The grey lines in the left panels show run LV2 (no annual cycle of α_3, just the high frequency variability with $P_2 = 0.2$ years). The grey lines in the right panels show the result of frequency dependent nudging to the climatology (run LV4). Parameter settings for all runs are given in Table 1.

Table 1
Description and parameters of the Lotka–Volterra (LV) model runs. All runs were 40 years long, $\alpha_1 = \alpha_2 = \alpha_4 = \alpha_5 = 50$, and $\alpha_{30} = \alpha_6 = 2$. The initial condition for run LV0 was $(x_{10} + 0.1, x_{20})$ where (x_{10}, x_{20}) is the interior fixed point. For runs LV1 to LV4 the initial condition was (x_{10}, x_{20}) and the period of the high frequency forcing was $P_2 = 0.2$. As detailed below, the runs differ only in the time variation of α_3 (see Eq. (8)) and the form of nudging.

Code	Description	α_{31}	α_{32}
LV0	No time variation of α_3	0	0
LV1	Full time variation of α_3	.2	.4
LV2	No annual cycle of α_3	0	.4
LV3	Conventional nudging of LV2	0	.4
LV4	Frequency dependent nudging of LV2	0	.4

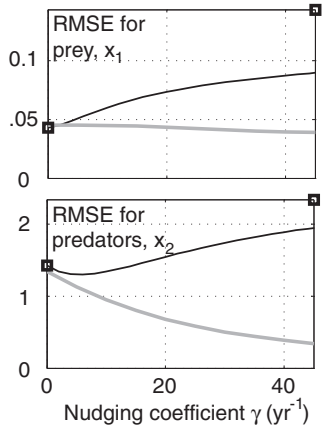

Fig. 3. RMSE of the LV model for conventional (black lines) and frequency dependent (grey lines) nudging as a function of γ. RMSEs were calculated over the last year of 40 year model runs. The squares at $\gamma = 0$ and $\gamma = 45$ show the RMSE for predictions based on the simplified model and climatology, respectively.

(see Table 1) are identical except for the amplitude of the annual cycle of α_3 and the form of nudging. Run LV1 includes the full time variation of carrying capacity and is not nudged. We will treat LV1 as the *complete model* and sample it to generate *observations* (see black lines of Figs. 1 and 2). Run LV2 is identical to LV1 except that $\alpha_{31} = 0$ leading to a seasonally biased simulation. We will treat LV2 as the *simplified model* (see gray lines in the left panels of Fig. 2). Runs LV3 and LV4 are identical to LV2 except that they are nudged to the mean and annual cycle of LV1 using conventional and frequency dependent nudging, respectively.

We implemented the climatological bandpass filter denoted by the angle brackets in (6) using a third-order Butterworth filter defined in state space form. The cutoff frequency of the lowpass filter is 1/6 cycle per year and the passband of the annual filter is [0.95, 1.05] cycle per year. The state space model for this filter was then combined with the predator–prey model by augmenting the predator–prey state vector, similar to the approach used by Thompson et al. (2006). The solution of (6) was then calculated numerically using an explicit Runge–Kutta scheme (ode45 routine in Matlab). Note that, in accordance with most theoretical studies of predator–prey interactions, we formulated (6) in continuous time. One consequence is that the nudging parameter in (6) is measured in units of reciprocal time and is limited solely by the constraint that it is nonnegative. Later in this study we will introduce a discrete time formulation for a more realistic biogeochemical model. For this discrete time formulation the nudging parameter will be dimensionless and constrained to lie between 0 and 1 (see Section 4).

One of the difficulties in implementing nudging is the specification of an appropriate nudging coefficient γ. The approach used here is to perform multiple runs of LV3 and LV4 with a range of γ and select the one with the lowest mean square error (MSE) relative to the complete run. For a trial γ to be considered valid we simply checked that the model reached a periodic steady state by the end of the run. With a stability coefficient of $\delta = 2$, we were able to obtain periodic solutions for γ less than about 50 yr^{-1}.

The black lines in Fig. 3 show the root MSE for conventional nudging as a function of γ. For $\gamma = 0$ the nudged run equals LV1 (see gray lines in the left panels of Fig. 2). As γ increases, the conventionally nudged solution approaches the climatology (dashed lines, left panels of Fig. 2). Fig. 3 shows that conventional nudging does not improve the model solution for the prey regardless of which value of γ is chosen. For values of $\gamma < 15\ \mathrm{yr}^{-1}$ the solution

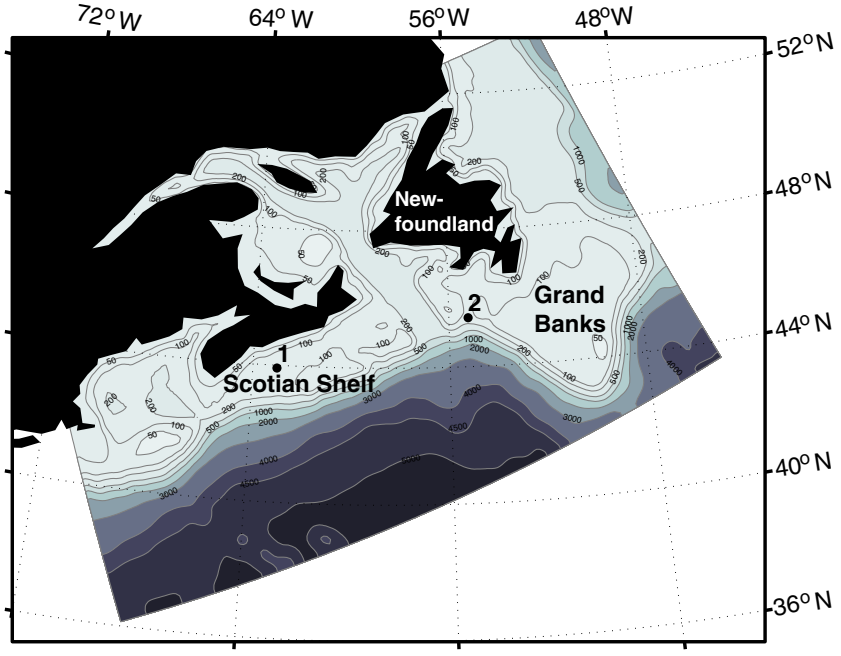

Fig. 4. Model domain and bathymetry of the complete 3D model of the northwest North Atlantic region (BO1). Synthetic observations were extracted for two locations: one on the Scotian Shelf (station 1) and one on the western tail of the Grand Banks (station 2).

for the predators improves only slightly. For $\gamma > 15$ yr^{-1} the solution degrades for the predators as well. Fig. 3 also shows that the MSE does not change monotonically with increasing γ. This is consistent with the complicated form of the transfer function for conventional nudging (see (3)).

The root MSE for frequency dependent nudging is shown by the gray lines in Fig. 3. For both x_1 and x_2 the MSE drops monotonically as $\gamma \to \infty$ and is well below the MSE for conventional nudging. Based on Fig. 3 we selected 45 yr^{-1} as the optimal γ value for frequency dependent nudging. The time variation of x_1 and x_2 for this choice of γ is shown by the gray lines in the right panels of Fig. 2. Frequency dependent nudging has clearly reduced the bias in the model state of LV2 in terms of the mean and annual cycle without suppressing the high frequency variability; the enhanced high-frequency variations of prey abundance when predator abundance is low has also been recovered.

The above set of experiments shows that frequency dependent nudging of a highly idealized, non-linear biological model in only two frequency bands can be effective, at least for the parameters given in Table 1.

4. Nudging a 1D biological ocean model

We now compare conventional and frequency dependent nudging using a more realistic, vertically resolved, biological ocean model configured for the continental shelf seas of the northwestern North Atlantic Ocean. The overall approach is identical to that used in the previous section. More specifically, a complete model generates observations and a simplified model generates seasonally biased results. Conventional and frequency dependent nudging are then used to suppress the seasonal bias in the state of the simplified model, and their impact on subseasonal variability is assessed by comparison with the observations. Again, the climatology used for nudging represents only the mean and the annual cycle (i.e. a sinusoid with period 1 year). We refer to these four models using codes that mimic those used previously for the LV model: BO1 and BO2 refer to the complete and simplified models, respectively, and BO3 and BO4 refer to the simplified model with conventional and frequency dependent nudging, respectively. However, in contrast to the LV models of the previous section the BO models are defined in discrete time.

4.1. Complete model and observations

The complete model (BO1) is a 3D, physical-biological model of the northwestern North Atlantic shelf seas and adjacent deep ocean (Fig. 4). It is an implementation of the Regional Ocean Modeling System (ROMS, http://myroms.org, (Haidvogel et al., 2008)) coupled to the biological model of Fennel et al. (2006, 2008). The model is described in detail by Bianucci et al. (submitted for

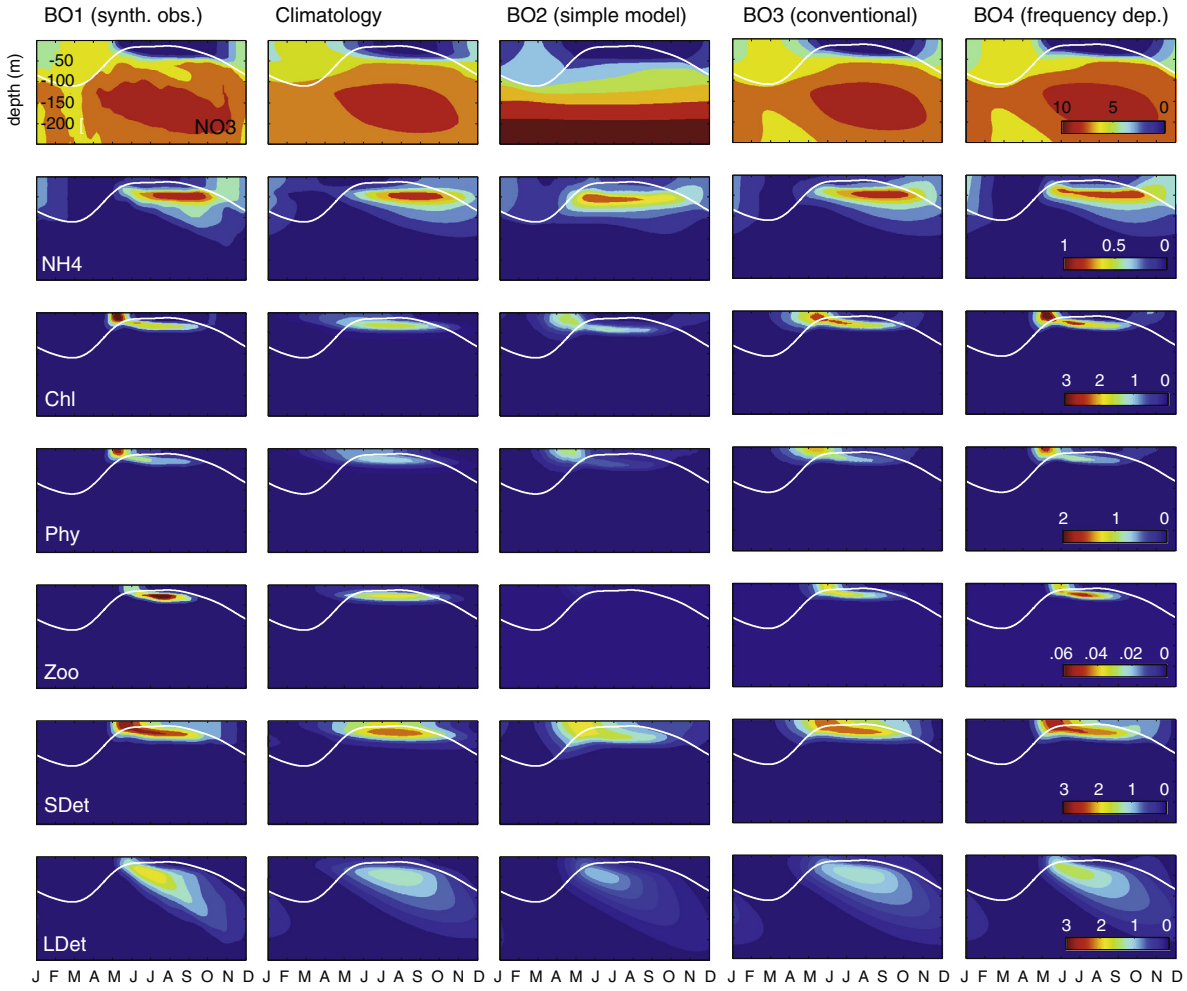

Fig. 5. From left to right: synthetic observations generated by the complete model BO1, climatology derived from BO1 consisting only of the mean and a sinusoid with period 1 year, simulation of the simple model (BO2), simulation of the simple model with conventional nudging (BO3) and simulation of the simple model with frequency dependent nudging (BO4). All plots are for Station 1 (see Fig. 4 for location). White lines show the evolution of the mixed layer depth. All concentrations are in mmol N m^{-3} except for chlorophyll which is in mg chl m^{-3}.

publication). The model domain (Fig. 4) includes the Grand Banks, the Gulf of St. Lawrence, the Scotian Shelf and the Gulf of Maine and is nested within the larger-scale physical model of Urrego Blanco and Sheng (2012). The model's horizontal resolution is ~10 km with 30 sigma-layers that are chosen to give higher resolution near the surface. The model is forced with atmospheric reanalysis fields for wind, heat and freshwater fluxes (Large and Yeager, 2004) and incoming solar radiation is prescribed using shortwave radiation estimates from the National Center for Environmental Prediction (NCEP, http://www.ncep.noaa.gov/).

The biological model component is a relatively simple representation of the marine nitrogen cycle that includes two species of dissolved inorganic nitrogen (nitrate, *NO*3, and ammonium, *NH*4), one functional phytoplankton group, *Phy*, chlorophyll, *Chl*, as a separate state variable to allow for photoacclimation, one functional zooplankton group, *Zoo*, and two pools of detritus representing large, fast-sinking particles, *LDet*, and suspended, small particles, *SDet*. The main processes described in the model are (1) temperature-, light- and nutrient-dependent phytoplankton growth with ammonium inhibition of nitrate uptake, (2) zooplankton grazing represented by a Holling-type III parameterization, (3) aggregation of phytoplankton and small detritus to fast sinking large detritus, (4) photoacclimation (i.e. a variable ratio between phytoplankton and chlorophyll), (5) linear rates of phytoplankton mortality, zooplankton basal metabolism, and detritus remineralization, (6) a second order zooplankton mortality, (7) light-dependent nitrification (i.e. oxidation of ammonium to nitrate), and (8) vertical sinking of phytoplankton and detritus. For further details on biological model justification, equations and parameters we refer the reader to Fennel et al. (2006, 2008).

The model was run from 1 January 1999 to 31 December 2004 and the model state was saved every five days. All seven biological state variables and temperature and salinity were extracted from the five-year simulation for the two stations shown in Fig. 4. The first year was discarded as spin-up. Observations representing the truth were created by regressing the times series of the remaining four years on the mean, a sinusoid with period 1 year representing the annual cycle and the next 14 harmonics (for each layer at both stations). The climatology was then obtained by retaining only the mean and annual cycle (identical to the approach used in the LV experiments). Note that again the simple model is going to be nudged toward the climatology (i.e. mean and annual cycle only), while the results from the nudging experiments will be compared against the observations (i.e. the mean, annual cycle and its 14 harmonics). The observations and climatology for Stations 1 and 2 are shown in the two left-most columns of panels in Figs. 5 and 6, respectively.

4.2. Simple model and biased simulation

The simple model (BO2) is 1D in space (only representing the vertical direction) and thus ignores processes like horizontal advection. It has a highly simplified physical component (described in more detail below) with a uniform vertical grid (spacing of 5 m) and uses the same biological model as BO1. BO2 has been configured for the two locations shown in Fig. 4.

The physical component of BO2 vertically diffuses the biological variables subject to no-flux boundary conditions at the surface and bottom. The Crank–Nicolson scheme is used to integrate the diffusion equation forward in time with a time step of six hours. Vertical diffusivities are defined in terms of a temporally varying mixed layer that was based on visual estimates of the mixed layer depth from individual temperature profiles from the 3D model. The mixed layer depth for Station 1 is shown in Fig. 5. The vertical

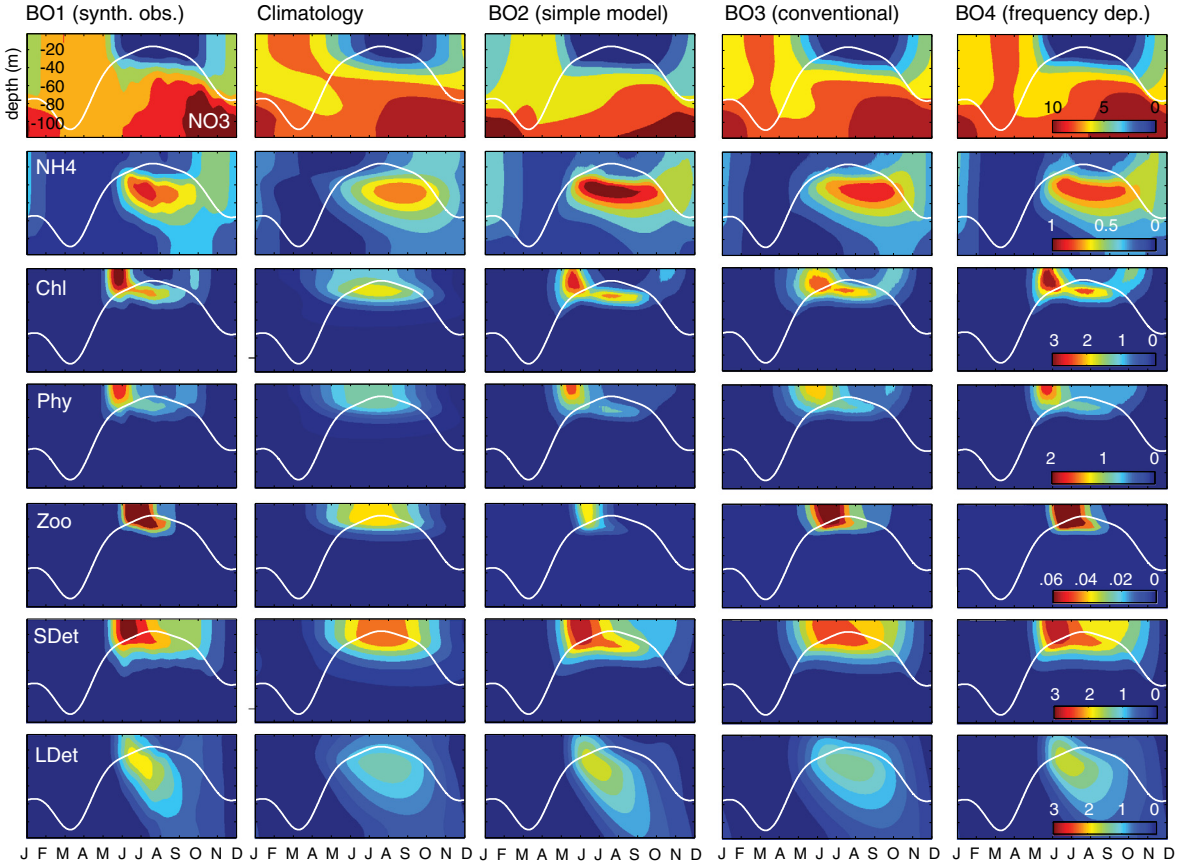

Fig. 6. As Fig. 5 but for Station 2.

diffusivities in and below the mixed layer (referred to as v_1 and v_2, respectively) are constant in the vertical.

Initial simulations of BO2 showed that the choice of vertical diffusivities is critical in determining the vertical structure of the biological variables. We evaluated a range of diffusivities by using the physical component of BO2 to simulate temperature. Comparisons with the temperature time series of BO1 indicated that vertical diffusivities should vary seasonally, presumably because of changes in stratification and wind mixing. A simple and effective way of capturing this seasonality is to allow the diffusivities to vary with the time-varying mixed layer depth (h) as follows:

$$v_{1,2} = (1-q)v_{1,2}^{winter} + qv_{1,2}^{summer} \tag{9}$$

where q is the normalized mixed layer depth

$$q = \frac{h_{max} - h}{h_{max} - h_{min}} \tag{10}$$

Based on these comparisons we chose $v_1^{winter} = 70\,\mathrm{m^2\,day^{-1}}$, $v_1^{summer} = 10\,\mathrm{m^2\,day^{-1}}$ for the upper layer. These diffusivities were reduced by a factor of 10 for the lower layer.

As stated above, the biological model in BO2 is the same as in the complete model BO1. In addition to the prescribed mixed layer variation the biological model is forced by temperature time series from BO1 and incoming shortwave radiation that drives phytoplankton growth but does not affect mixing. The shortwave radiation for BO2 is based on daily integrated values from the NCEP data set (see above), interpolated to the horizontal position of the station under consideration. These daily values do not include a diurnal cycle, while ROMS imposes a diurnal cycle internally within its biological module by redistributing the daily integral of incoming solar radiation according to the theoretical diurnal cycle determined by astronomical formulae. The time step of the ROMS model is about a minute, which ensures that the diurnal cycle is resolved very well in BO1. BO2 has a time step of six hours, which is insufficient to resolve the diurnal variations. In an attempt to capture the main features of the diurnal cycle in BO2 we simply designated two time steps as night (setting incoming solar radiation to zero) and distributed the daily-integrated solar radiation equally over the other two time steps (designated as day). This ensures that BO2 receives the same daily integral of solar radiation as BO1.

The biological variables of BO2 are integrated forward in discrete time by first applying the vertical mixing step (Crank–Nicolson scheme) and then a biological update step (Euler forward scheme). BO2 was integrated for 15 years and had reached a periodic steady state by the end of the run. The final year is shown in Figs. 5 and 6 for Stations 1 and 2, respectively. There are clearly significant differences between the last year of BO2 and the observations from BO1: at Station 1 the nitrate concentration at depth is too high; at both stations the zooplankton concentration is too low; the peak phytoplankton concentration during the spring bloom is too low, particularly at Station 1; at both stations the concentration of detritus is too low. Thus, BO2 is a biased model and

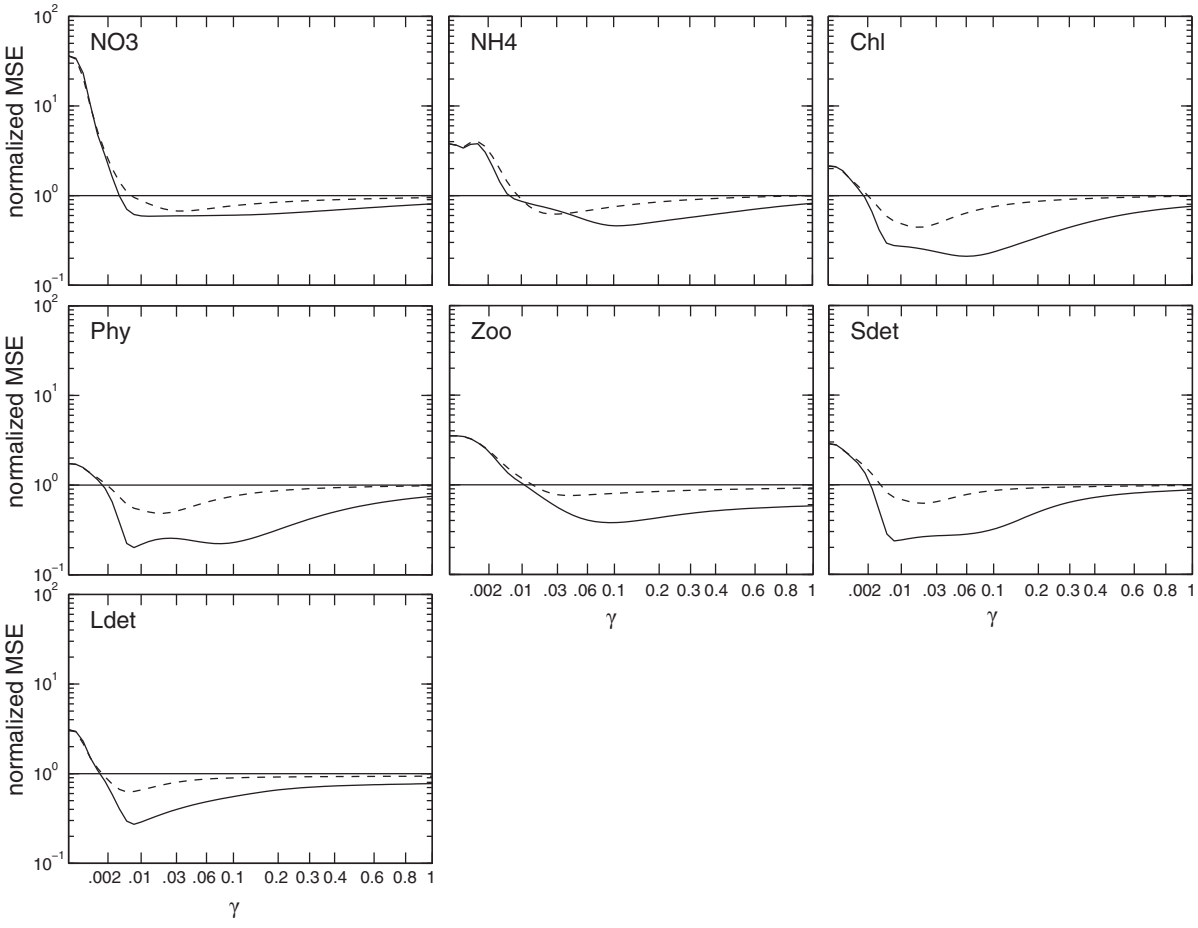

Fig. 7. Mean square error (MSE) between observations (from BO1) and nudged simulations (BO3 and BO4) for Station 1. MSEs are normalized by the MSE between observations and climatology and plotted as a function of nudging coefficient γ. All MSEs were calculated over the top 100 m for the last year of model integration and are shown for conventional nudging by dashed lines and for frequency dependent nudging by solid lines. The x-plotting positions have been stretched to vary as γ^3 in order to expand the plot near $\gamma = 0$.

represents a good test case for assessing the effects of different nudging schemes.

4.3. Effect of nudging

We now nudge the simplified model using the climatology consisting only of the mean and annual cycle of BO1. Conventional and frequency dependent nudging were implemented in BO3 and BO4 using nudging coefficients γ that have been normalized by the model time step. The nudging coefficient is therefore nondimensional and ranges between 0 (no nudging) and 1 (direct insertion of the climatology into the model).

The frequency dependent nudging was implemented as in Eq. (6) except that (i) the model is now formulated in discrete time, and (ii) the nudging term added to the updated model state is of the form $\gamma[(1-\delta)\langle c_n - x_n\rangle + \delta(c_n - x_n)]$ where $c_n - x_n$ is the difference between the climatology and updated model state at time n. The bandpass filter applied to the differences is recursive in form and is described by Thompson et al. (2006). In short, the filtering corresponds to regressing $c_n - x_n$ on a constant and annual cycle using a sliding window and then estimating the model state at the present time using the fitted regression model. The effective width of the sliding window and the bandwith of the filter were set by choosing $\kappa = \frac{1}{4}$ yr^{-1} (see Thompson et al., 2006 for further discussion of this parameter).

We took the same approach to choosing the nudging coefficient as with the LV model, that is, we performed multiple nudging runs with γ ranging between 0 and 1. For each run we calculated the MSE between the observations from the complete model (BO1) and the last year of the nudged runs (BO3 and BO4). The dependence of MSE on γ is shown in Fig. 7 for Station 1. Clearly, nudging improves the fit of the simple model for all variables. The improvement is markedly better for frequency dependent nudging, especially for chlorophyll, phytoplankton, zooplankton and detritus. The improvement due to nudging is often sustained over larger ranges of γ for the frequency dependent nudging. The γ values of minimum MSE are not identical for all variables, hence there is no obvious choice of the optimal γ. However, it is easier to choose an optimal value for frequency dependent nudging because of the broad minima in MSE. We chose $\gamma = 0.020$ and 0.025 for conventional and frequency dependent nudging, respectively.

Nudging improves the results of the simple model for both conventional and frequency dependent nudging (Fig. 5). At Station 1 the most obvious difference between the observations (BO1) and the simple model (BO2) is in the vertical structure of the nitrate distribution (nitrate concentrations between 50 and 100 m depth

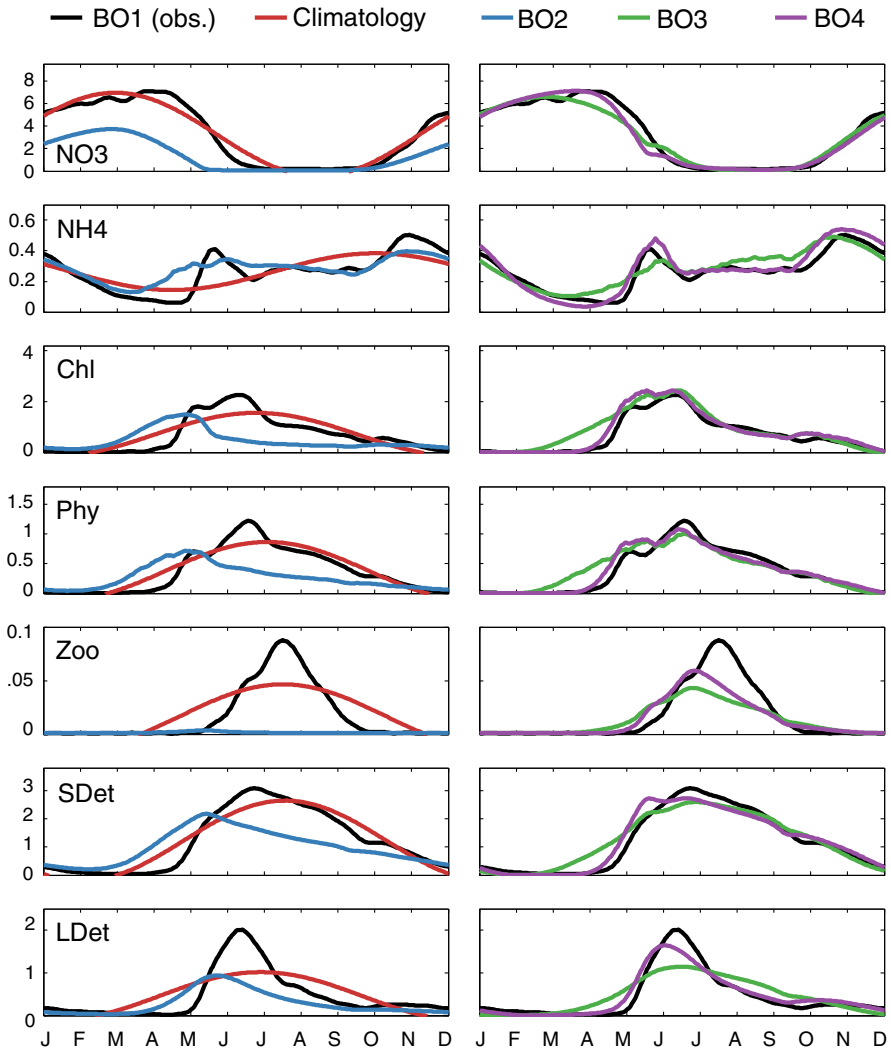

Fig. 8. Time series of biological variables at 30 m depth for the observations (BO1), the climatology (i.e. mean and sinusoid with period 1 year), the simple model (BO2), the simple model nudged conventionally to climatology (BO3), and the simple model with frequency dependent nudging (BO4) for Station 1. All concentrations are in mmol N m^{-3} except for chlorophyll which is in mg chl m^{-3}.

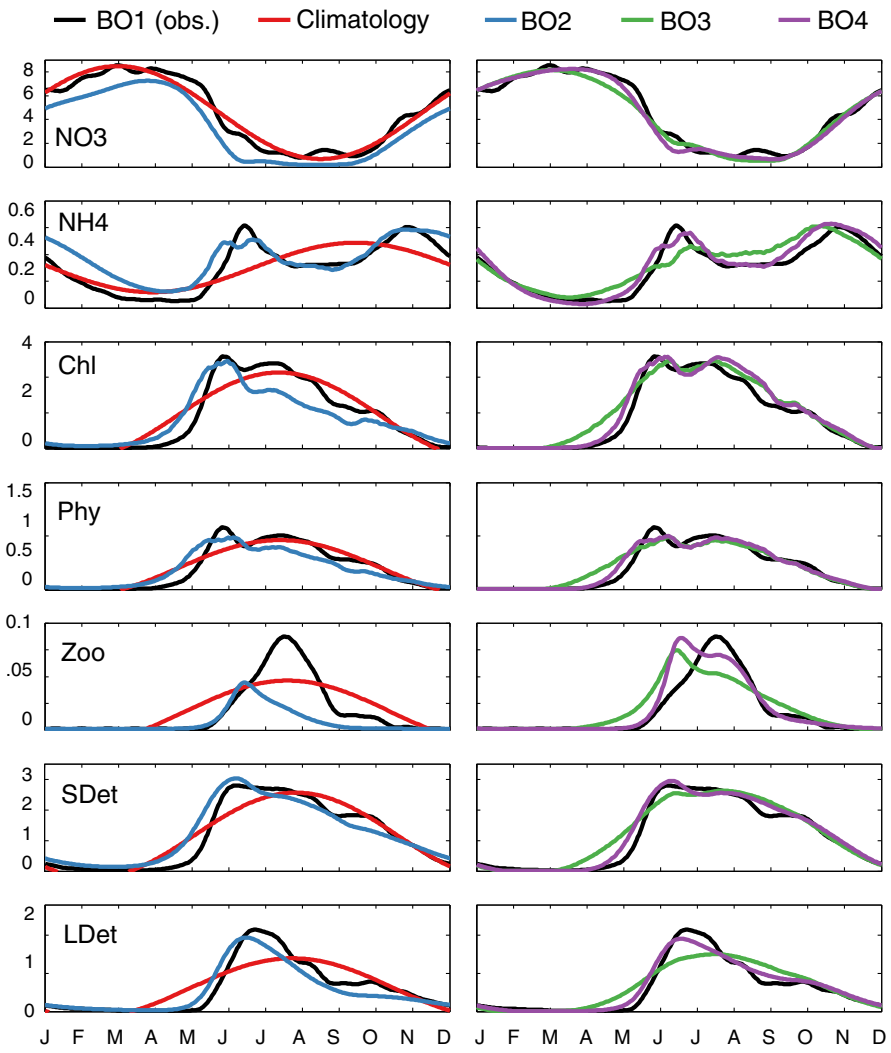

Fig. 9. As Fig. 8 but for Station 2.

Table 2
Quantitative comparison of the performance of conventional and frequency dependent nudging at the two stations shown in Fig. 4. Given are the mean square errors (MSEs) between the observations (BO1) and simulations by the simplified model without nudging (BO2), with conventional nudging (BO3) and with frequency dependent nudging (BO4). The MSEs are calculated for the last year of model integration over the whole water column and have been normalized by the MSE between observations and climatology. The values of γ used for the three runs are also given. The frequency dependent run used a stability coefficient of $\delta = 0.1$.

Run	γ	*NO3*	*NH4*	*Chl*	*Phy*	*Zoo*	*SDet*	*LDet*
Station 1								
BO2	0	24.48	4.30	1.60	1.42	3.29	2.26	1.61
BO3	0.020	0.74	0.65	0.45	0.49	0.82	0.62	0.78
BO4	0.025	0.62	0.70	0.24	0.25	0.58	0.27	0.41
Station 2								
BO2	0	3.31	2.16	0.56	0.66	1.40	0.86	0.33
BO3	0.020	0.71	0.93	0.50	0.51	0.33	0.70	0.73
BO4	0.025	0.38	0.51	0.22	0.20	0.11	0.24	0.18

are much lower in BO2 than BO1; conversely, below 200 m nitrate concentrations are much higher in BO2 than BO1). The poor representation of the vertical nitrate distribution in BO2 is a major factor in for the overall deterioration of results in BO2 at station 1. Both nudging schemes (BO3 and BO4) dramatically improve the vertical nitrate distribution (essentially by adding nitrate between 50 and 100 m depth and removing nitrate below 200 m). This results in an increased and more realistic supply of nitrate to the mixed layer in winter. The only difference between the conventional and frequency dependent nudging cases is that surface nutrients disappear more quickly during spring in the latter case. The variable that is least affected by nudging is ammonium, which is not surprising given that ammonium distributions are very similar between observations, climatology and simple model. Chlorophyll and phytoplankton, both significantly underestimated in the simple model, have increased spring maxima with conventional nudging, but still underestimate the peak of the spring bloom. With frequency dependent nudging, chlorophyll and phytoplankton peaks are much closer to the observations. Subsurface, summer maxima of zooplankton, which essentially disappeared in the simple model without nudging, are present in the conventionally nudged model and increase further in the frequency dependent nudging case, leading to more realistic results in the latter. Small and large detritus respond to nudging in a similar way (conventional nudging does improve the results, but with a more pronounced improvement with frequency dependent nudging).

In Fig. 8 we show time series of all variables at 30 m depth. This figure illustrates the smoothness of the climatology used for nudging, and how the simple model with frequency dependent nudging is better able to reproduce concentration maxima (e.g. in ammonium, zooplankton and large detritus) and periods of rapid increase/decrease (e.g. the spring drawdown of nitrate and spring

increase of ammonium, chlorophyll and phytoplankton) which are steeper with frequency dependent nudging.

At Station 2, which is much shallower than Station 1, the evolution and vertical structure of nitrate is better captured by the simple model than at Station 1, although supply during winter mixing is underestimated at this station as well (Fig. 6). Both nudging approaches improve this aspect of the simulation. The simple model overestimates subsurface ammonium concentrations in summer, slightly underestimates the spring maxima in chlorophyll and phytoplankton, and significantly underestimates zooplankton. The evolution of ammonium and zooplankton are significantly improved with both nudging approaches, but the improvements for chlorophyll and phytoplankton are much more obvious for frequency dependent nudging than conventional nudging. Time series plots (Fig. 9) again show how the simple model with frequency dependent nudging is better able to reproduce periods of rapid change such as the nitrate drawdown during spring and the associated increases in the other variables.

A quantitative assessment of conventional and frequency dependent nudging at the two stations is provided in Table 2. At Station 1, either form of nudging markedly improves the results compared to the model without nudging, often by significantly more than 50%. Frequency dependent nudging outperforms conventional nudging by improving the results by another 30 to 50% except for nitrate, which is improved by only 16%, and ammonium, which is slightly degraded when compared to the conventional nudging case. The slightly smaller improvement of ammonium at Station 1 is the only case where conventional nudging outperforms frequency dependent nudging. At Station 2, conventional nudging again improves the results compared to the un-nudged simulation (except for large detritus), however, the improvement is much less pronounced than at Station 1, especially for chlorophyll and phytoplankton. At this station, frequency dependent nudging leads to significant improvements of 46 to 65% compared to conventional nudging.

5. Summary and discussion

We explored and compared the application of conventional and frequency dependent nudging to biological models. First we applied the nudging methods to a simple predator–prey model (the LV model) and then to a 1D biogeochemical ocean model for the northwestern North Atlantic shelf seas (the BO model). Our approach was to first create observations from a complete model and compute a smooth climatology based on the mean and annual cycle (sinusoid with period 1 year) from these observations. We then simplified the model such that its results were biased and applied conventional and frequency dependent nudging using the climatology. Qualitative and quantitative comparisons between the observations and nudged model results showed that frequency dependent nudging outperformed conventional nudging in practically every case and better allowed the nonlinear models to recover much of the higher frequency variability. For the LV runs conventional nudging suppressed variability on sub-seasonal timescales and generally degraded results while frequency dependent nudging led to improvements. Our nudging experiments with the BO model showed that conventional nudging often improves biased results, but that frequency dependent nudging leads to further, significant improvements.

Several limitations should be noted however. First, our conclusions are limited to the cases studied here, which use synthetically generated observations. Second, the nudging methods described here only reduce biases in the simulated model state, not the model itself. Thus, these techniques are no substitute for fixing errors in the models structure, parameterizations or forcing that can be fixed. Nevertheless, some bias errors will likely remain in realistic models and techniques for online bias reduction will continue to be a necessary procedure in operational forecasting and the generation of optimal hindcasts. We note however that the spatial and temporal structure of the applied nudges may be useful in identifying the cause of systematic model errors, e.g. erroneous vertical diffusivities would be indicated by nudges of opposite sign in the vertical direction.

Our experiments suggest that frequency dependent nudging is a promising technique for the reduction of biases in biogeochemical model states, although firm conclusions are necessarily limited to the cases we have studied here. As a next step the technique will be applied to a 3D biogeochemical model.

Acknowledgments

This work was supported by the Ocean Tracking Network Canada. We wish to thank two anonymous reviewers for insightful comments that helped improve the manuscript.

References

Baek, S.-J., Hunt, B.R., Kalnay, E., Ott, E., Szunyogh, I., 2006. Local ensemble Kalman filtering in the presence of model bias. Tellus 58A, 293–306.

Bagniewski, W., Fennel, K., Perry, M.J., D'Asaro, E.A., 2011. Optimizing models of the North Atlantic spring bloom using physical, chemical and bio-optical observations from a Lagrangian float. Biogeosciences 8, 1291–1307.

Berryman, A.A., 1992. The origins and evolution of predator–prey theory. Ecology 73, 1530–1535.

Bianucci, L., Fennel, K., Mattern, P., submitted to Ocean Modelling. Optimizing phytoplankton dynamics in a 3D regional model of the northwest North Atlantic, submitted for publication.

Chepurin, G.A., Carton, J.A., Dee, D., 2005. Forecast model bias correction in ocean data assimilation. Mon. Weather Rev. 133, 1328–1342.

Chesson, P., 1990. MacArthur's consumer-resource model. Theor. Populat. Biol. 37, 26–38.

Dee, D.P., Todling, R., 2000. Data assimilation in the presence of forecast bias: the GEOS moisture analysis. Mon. Weather Rev. 128, 3268–3282.

Doney, S.C., Lima, I.D., Moore, J.K., Lindsay, K., Behrenfeld, M.J., Westberry, T.K., Mahowald, N.M., Glover, D.M., Takahashi, T., 2009. Skill metrics for confronting global upper ocean ecosystem-biogeochemistry models against field and remote sensing data. J. Mar. Syst. 76, 95–112.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., Liebert, J., 2012. Should we apply bias correction to global and regional climate model data? Hydrol. Earth Syst. Sci. 16, 3391–3404.

Fennel, K., Abbott, M.R., Spitz, Y.H., Richman, J.G., Nelson, D.M., 2003. Modeling controls of phytoplankton production in the southwest Pacific sector of the Southern Ocean. Deep-Sea Res. II 50, 769–798.

Fennel, K., Wilkin, J., Levin, J., Moisan, J., O'Reilly, J., Haidvogel, D., 2006. Nitrogen cycling in the Middle Atlantic Bight: Results from a three-dimensional model and implications for the North Atlantic nitrogen budget. Global Biogeochem. Cycles 20, GB3007, http://dx.doi.org/10.1029/2005GB002456.

Fennel, K., Wilkin, J., Previdi, M., Najjar, R., 2008. Denitrification effects on air-sea CO_2 flux in the coastal ocean: Simulations for the northwest North Atlantic. Geophys. Res. Lett. 35, L24608. http://dx.doi.org/10.1029/2008GL036147.

Haidvogel, D., Arango, H., Budgell, W., Cornuelle, B., Curchitser, E., Di Lorenzo, E., Fennel, K., Geyer, W., Hermann, A., Lanerolle, L., et al., 2008. Ocean forecasting in terrain-following coordinates: formulation and skill assessment of the regional ocean modeling system. J. Comput. Phys. 227, 3595–3624.

Keppenne, C.L., Rienecker, M.M., Kurkowski, N.P., Adamec, D.A., 2005. Ensemble Kalman filter assimilation of temperature and altimeter data with bias correction and application to seasonal prediction. Nonlinear Process. Geophys. 12, 491–503.

Large, W., Yeager, S., 2004. Diurnal to Decadal Global Forcing for Ocean and Sea-Ice Models: The Data Sets and Flux Climatologies. National Center for Atmospheric Research, Boulder, Colorado, USA.

Lehmann, M.K., Fennel, K., He, R., 2009. Statistical validation of a 3-D bio-physical model of the western North Atlantic. Biogeosciences 6, 1961–1974.

MacArthur, R.H., 1970. Species packing and competitive equilibria for many species. Theor. Populat. Biol. 1, 1–11.

Marchal, O., Stocker, T.O., Joos, F., 1998. A latitude-depth, circulation-biogeochemical model for paleoclimate studies. Develop. Sens. Tellus 50B, 290–316.

May, R.M., 1973. Stability and Complexity in Model Ecosystems. Princeton University Press.

Najjar, R.G., Sarmiento, J.L., Toggweiler, J.R., 1992. Downward transport and fate of organic matter in the ocean: simulations with a general circulation model. Global Biogeochem. Cycles 6, 45–76.

Nerger, L., Gregg, W.W., 2007. Assimilation of SeaWiFS data into a global ocean-biogeochemical model using a local SEIK filter. J. Mar. Syst. 68, 237–254.

Nerger, L., Gregg, W.W., 2008. Improving assimilation of SeaWiFS data by the application of bias correction with a local SEIK filter. J. Mar. Syst. 73, 87–102.
Priestley, M.B., 1981. Spectral Analysis and Time Series. Academic Press.
Randall, D.A., Wood, R.A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R.J., Sumi, A., Taylor, K.E., 2007. Climate models and their evaluation. In: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (Eds.), Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
Stacey, M., Shore, J., Wright, D., Thompson, K., 2006. Modeling events of sea-surface variability using spectral nudging eddy permitting model of the Northeast Pacific Ocean. J. Geophys. Res. 111, C06037. http://dx.doi.org/10.1029/2005JC003278.
Thompson, K.R., Wright, D.G., Lu, Y., Demirov, E., 2006. A simple method for reducing seasonal bias and drift in eddy resolving ocean models. Ocean Model. 13, 109–125.
Thompson, K.R., Ohashi, K., Sheng, J., Bobanovic, J., Ou, J., 2007. Suppressing bias and drift of coastal circulation models through the assimilation of seasonal climatologies of temperature and salinity. Cont. Shelf Res. 27, 1303–1316.
Urrego Blanco, J.U., Sheng, J., 2012. Numerical investigation of inter annual variability of circulation and hydrography over the eastern. Canadian Shelf. Atmos. Ocean 50, 277–300.
While, J., Haines, K., Smith, G., 2010. A nutrient increment method for reducing bias in global biogeochemical models. J. Geophys. Res. 115, C10036. http://dx.doi.org/10.1029/2010JC006142.
WMO WWRP 2009-1, 2008. Recommendations for the Verification and Intercomparison of QPFs and PQPFs from Operational NWP Models. WMO/TD – No. 1485, World Meteorological Organization.
Woodgate, R.A., Killworth, P.D., 1997. The effects of assimilation on the physics of an ocean model. Part 1: theoretical model and barotropic results. J. Atmos. Ocean Technol. 14, 897–909.
Zhu, J., Demirov, E., Dupont, F., Wright, D., 2010. Eddy-permitting simulations of the sub-polar North Atlantic: impact of the model bias on water mass properties and circulation. Ocean Dyn. 60, 1177–1192.